Evolution of Plastic Neurocontrollers for Situated Agents

Evolution of Plastic Neurocontrollers for Situated Agents

Dario Floreano
Laboratory of Cognitive Technology
AREA Science Park
Trieste, Italy
dario@psicoiris.area.trieste.it

Francesco Mondada
Laboratory of Microcomputing
Swiss Federal Institute of Technology
Lausanne, Switzerland
mondada@di.epfl.ch

Abstract

In this paper we investigate a novel approach to the evolutionary development of autonomous situated agents based on the assumption that the neural mechanisms underlying ontogenetic learning are themselves developed and shaped by the evolutionary process. A genetic algorithm is used to evolve neural structures that can be continuously modified during life according to the mechanisms specified in the genotype. The evolutionary process is carried out on a real mobile robot. The analysis of one of the best evolved individuals shows rapid development of stable behavior mediated by fast-changing synapses which are dynamically stable.

1 Introduction

Adaptation to the environment takes place at multiple levels and time-scales: it ranges from the long-term dynamics of *phylogenetic evolution* to the fast process of *ontogenetic learning*. At the neural level, these adaptation modalities, that are well-timed and co-ordinated, are responsible for the development of robust and complex control systems that display the ability to self-regulate their own behavior and keep the organism alive.

Behaviors which are inherited at birth and cannot be modified by experience are called *innate behaviors* or *instincts*[1] [11]. To this extent, most of the experiments on *animats* which involve some form of artificial evolution have been concerned with the evolution of innate behaviors. However, there are a few experiments which combine learning and evolution making use of carefully designed architectures and traditional supervised learning algorithms (e.g., see [13, 1, 5]).

In this paper we investigate a different approach based on the assumption that the neural mechanisms underlying ontogenetic learning are themselves developed and shaped by the evolutionary process (see section 5 for biological considerations). A genetic algorithm is used to evolve neural structures that can be continuously modified during life according to the mechanisms specified in the genotype. Each decoded network is downloaded into a mobile robot which is let free to interact with the environment while its fitness is automatically computed and stored away for selective reproduction.

[1] The term *instinct* has undergone several re-definitions which have stressed the influence of experience and maturation on the final behavior. Here we assume the definition given by Darwin, whereby instincts are the product of natural selection and inheritance.

2 Experimental setup and task description

The experimental setup employed in these experiments is identical to that already described in [4], which we summarize below. We used the miniature mobile robot Khepera [12] which has a circular shape with a diameter of 55 mm, a height of 30 mm, and a weight of 70 g. Khepera is supported by two wheels and two small Teflon balls. The two wheels are controlled by two DC motors equipped with incremental encoders (12 pulses per mm of advancement of the robot) and can move in both directions. Each motor controller sets the speed of its own wheel according to a continuous value between -0.5 and +0.5, where 0.0 means no rotation, -0.5 means maximum rotation speed in one direction (set to 80 mm/s) and 0.5 means maximum rotation speed in the opposite direction. Each of the eight Infra-red proximity sensors, six positioned facing one direction of motion and two the opposite direction, returns a continuous value between 0 and 1 that signals the distance of an object from that sensor (the closer the object, the higher the value returned). In our environment the maximum detection range was approximately 4 cm. The robot was provided with a small positioning device which detected light beams emitted by a laser device placed on the top of the environment and computed the robot absolute position (see [5] for further details). This information was used only for behavior analysis and was not passed to the neural controller. Khepera was attached via a serial port to a Sun SparcStation by means of a lightweight aerial cable and specially designed rotating contacts.

The robot was put in an environment consisting of a

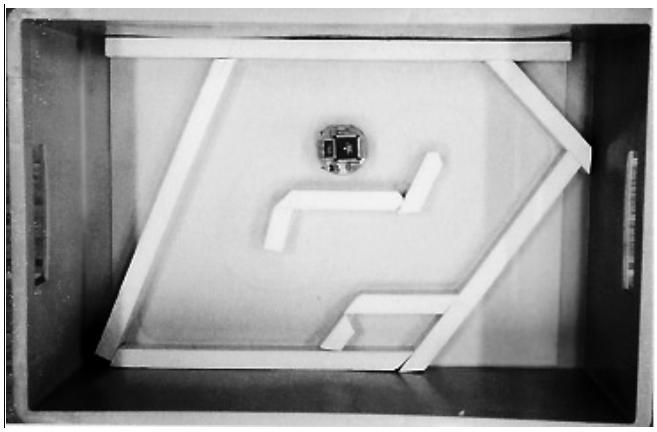

Figure 1: Bird-view of the environment with the robot.

circular corridor whose external size was approx. 60x50 cm large (Figure 1). The walls were made of light-blue polystyrene and the floor was made of gray paperboard. The robot could sense the walls with the IR proximity sensors. The environment was within a portable box positioned in a room always illuminated from above by a 60-watt bulb light.

The genetic operators, the decoding routines from genotypes to phenotypes, and the neural network dynamics were managed by the workstation CPU. Each individual of a population was in turn decoded into the corresponding neural network, the input nodes connected to the robot sensors, the output nodes to the motors (output unit activation was transformed in the range ± 0.5 before passing it to the motor), and the robot was let free to move for 24 s (80 motor actions) while its fitness Φ was automatically recorded and accumulated. Each sensory-motor loop lasted 300 ms (to/from communications between the robot and the workstation lasted approximately 60 ms) during which the wheel speed was kept constant. Between one individual and the next, a pair of random velocities was applied to the wheels for 5 seconds: this procedure was aimed at limiting the artifactual inheritance of particular locations between adjacent individuals in the population.

The fitness function Φ was designed to evolve obstacle avoidance and straight navigation behaviors, as in [4]

$$\Phi = V\left(1-\sqrt{\Delta v}\right)(1-i) \qquad (1)$$

$$\begin{aligned} 0 &\leq V &&\leq 1 \\ 0 &\leq \Delta v &&\leq 1 \\ 0 &\leq i &&\leq 1 \end{aligned}$$

where V is a measure of the average rotation speed in absolute value of the two wheels, $\Delta v = (v_{left} + 0.5) - (v_{right} + 0.5)$ is the absolute value of the difference between the speed of the wheels transformed into positive values, and i is the activation value of the proximity sensor with the highest activity. Φ was newly computed every 300 ms, accumulated during the life of the individual, and finally normalized by the number of actions. Since the robot has a circular shape and the wheels can rotate in both directions, this function has a symmetric surface with two equal maxima, each corresponding to one motion direction.

3 Genetic Encoding and Network Dynamics

A simple genetic algorithm (as described in [6]) with linear fitness scaling, roulette-wheel selection, one-point crossover, and bit-substitution mutation was employed (see details in the Appendix) to evolve binary chromosomes which encoded a set of parameters describing synapses properties and learning rules (see section 5 for motivations). Every time a phenotype was created, its synapses were intialized to small random values and could change their strength during life; final strengths were *not* coded back into the chromosome. Thus, each decoded neural network changed its own synaptic strength configuration according to the genetic instructions and without external supervision while the robot interacted with its own environment.

Each neural network had only three neurons — one hidden neuron and two motor neurons, each receiving synaptic connections from all the eight IR sensors and from the hidden neuron (Figure 2); this architecture could not be modified in the experiments described here.

Synaptic connections could have a *driving* or a *modulatory* effect on the postsynaptic neuron; afferent signals were combined in a two-component activation function [14] which gave an output between 0 and 1 (Figure 3). Driving signals determined whether the unit activity was below or above 0.5 (which, when transformed into the range ± 0.5 for motor control, was the point of inversion of wheel rotation), whereas modulatory signals could enhance or dampen the unit response, but could not change the direction of wheel rotation.

Synapses were individually coded on the chromosome. Each synapse was described by a set of four properties: whether it is driving or modulatory (1 bit), whether it is excitatory or inhibitory (1 bit), its learning rule (2 bits), and its learning rate (2 bits).

Each individual synapse could change its strength according to one of four basic Hebbian learning rules: pure Hebbian, postsynaptic, presynaptic, and covariance (see [20]). We have slightly modified each of these rules in order to satisfy the following constraint. Synaptic strength cannot grow indefinitely, but is intrinsically bound in the range $[0.0, 1.0]$ by means of a self-limiting mechanism which depends on the current synaptic strength; this so-

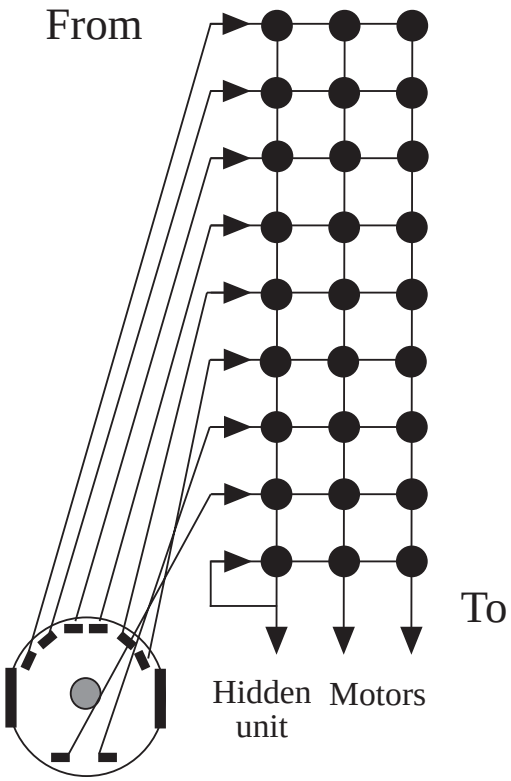

Figure 2: The architecture of the neural network employed. The black circles are the synapses; the circle in the middle of the robot body represents the hidden unit. The activations of the three units correspond — respectively — to the hidden units, the left motor, and the right motor.

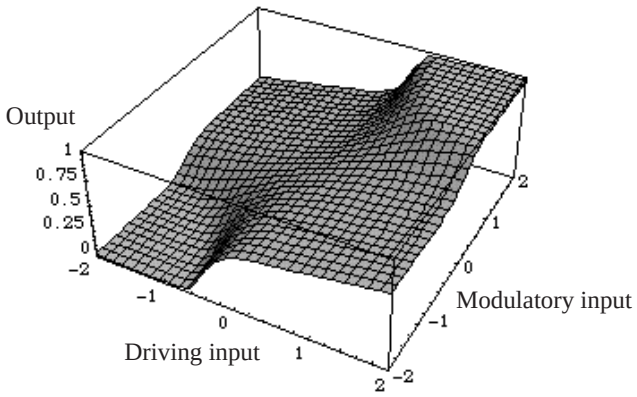

Figure 3: The activation function of internal nodes and motor nodes [14]: signals coming from driving connections and modulatory connections are separately integrated and passed to the network.

lution has the property of keeping the sign of the synapse unchanged, thus reducing the degrees of freedom of the network and putting more emphasis on the genetically evolved configuration of excitation and inhibition. The four types of synaptic change Δw are as follows (where x, y are respectively the pre- and postsynaptic activation and range between 0.0 and 1.0):

- The simplest learning mechanism is plain Hebb, whereby synapses can only be strengthened

$$\Delta w = (1 - w)\, xy \tag{2}$$

- The postsynaptic rule is similar to the plain Hebbian rule, but also decreases the synaptic efficacy when the postsynaptic unit is active and the presynaptic unit is not

$$\Delta w = w\,(-1 + x)\, y + (1 - w)\, xy \tag{3}$$

- Instead, in the presynaptic learning rule the decrement takes place when the presynaptic unit is active, but the postsynaptic unit is inactive

$$\Delta w = wx\,(-1 + y) + (1 - w)\, xy \tag{4}$$

- The covariance rule here takes the form of a synchronous- activation detector: if the presynaptic and postsynaptic activity levels differ by more than half the maximum node activation, the synaptic efficacy is reduced in proportion to that difference, otherwise it is increased in proportion to the difference

$$\Delta w = \begin{cases} (1 - w)\mathcal{F}(x, y) & \text{if } \mathcal{F}(x, y) > 0 \\ (w)\mathcal{F}(x, y) & \text{otherwise} \end{cases} \tag{5}$$

  where $\mathcal{F}(x, y) = \tanh(4(1 - |x - y|) - 2)$ is a measure of the difference between the presynaptic and postsynaptic activity. $\mathcal{F}(x, y) > 0$ if the difference is bigger or equal to 0.5 (half the maximum node activation) and $\mathcal{F}(x, y) < 0$ if the difference is smaller than 0.5.

As soon as the network is decoded and attached to the sensors and motors of the robot, synaptic weight values are intialized to small random values in the range $[0.0, 0.1]$ and are updated every 300 ms according to the following discrete-time equation

$$w_t = w_{t-1} + \eta \Delta w_t \tag{6}$$

where η is the learning rate, which can assume one of four values $\{0.0, 0.3, 0.7, 1.0\}$. If the learning rate is 0.0, that synapse will not change its strength during the life of the individual.

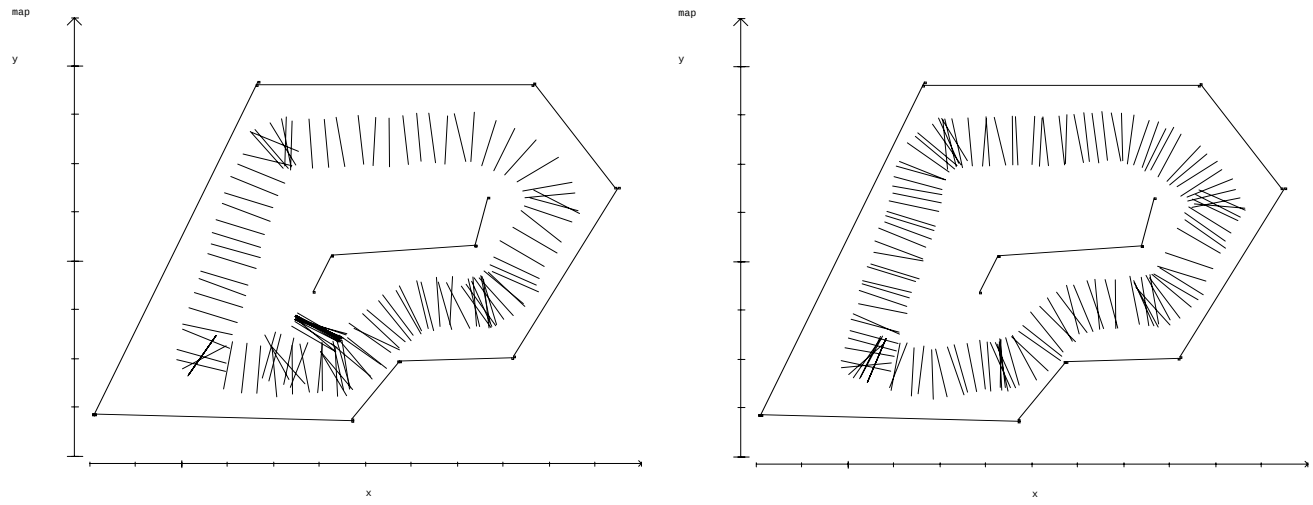

Figure 4: Trajectory of the robot that learns to navigate during life. Position data, visualized as bars representing the axis connecting the two wheels, were acquired with the laser positioning system every 100 ms. Data refer to the best individual of the last generation of one evolutionary run. Left: trajectory during the first lap (the robot starts in the lower portion of the environment and turns anti-clockwise). Right: trajectory at the second lap.

4 Results

Three different runs of this experiment were made. In all cases the best individual fitness reached a maximum value around the 50th generation ($\Phi = 0.23, \pm 0.09$). When compared to the results of the experiment reported in [4] where we evolved only synaptic strengths for the same task of obstacle avoidance and straight navigation, the fitness values recorded here displayed higher variation.

All the best neural networks of the last generation could control the robot in order to keep a straight trajectory while avoiding obstacles. The evolved behaviors resulted in smooth paths around the arena (Figure 4). This ability was developed by each individual neurocontroller during the first few sensory-motor loops, whatever the inital random values assigned to the synapses. In all the three runs the best individuals of the last generation moved in the direction where more IR sensors were placed, which provided a better sampling of the obstacles facing the robot.

The evolved neurocontrollers varied in the type of behavioral strategies and learning modalities both within a single population and across the three evolutionary runs. Here we present an analysis of the best individual of the last generation of one run and in section 5 we shall give some data on how this individual differed from the others.

The neural network was decoded, connected to the robot sensors and motors, the synaptic strengths were intialized to random values in the range $[0.0, 0.1]$, the robot was positioned facing a corner of the inner wall (Figure 4, left; initial position corresponds to the set of superimposed bars in the lower portion of the environment) and let free to move. During the first 2 s (6-7 synaptic updates) the robot adjusts its position alternating backward and forward motions until it finds a wall on its right side. This initial behavior is quite stereotypical: it is displayed for any starting position. Once the wall is found, the robot moves forward keeping the wall at a distance of 2 cm from its own right side; every second or third action, it slightly turns toward the wall and then continues on the previous direction. This sort of jerky behavior is gradually reduced when moving along a straight long wall (e.g., along the north and east walls). If the wall is slightly bent, the robot can still follow it without reducing speed, but when the walls form a convex angle smaller than 90 degrees (which means that most of the front IR sensors are active) the robot stops, backs and rotates to the right, and then moves forward again in the new direction. The robot has developed a sort of wall-following strategy. After one lap of the corridor, the path becomes smoother with less trajectory adjustements and more tuned to the geometric outline of the environment (Figure 4, right).

4.1 Internal Dynamics

The development of such behavior can be understood if one looks at the internal dynamics of the evolved network. Figure 5 plots the strengths of all the synapses in the network during the first 100 actions (sensory-motor loops) visualized in Figure 4 where the plots are laid out in the same format as the synapses in Figure 2. The

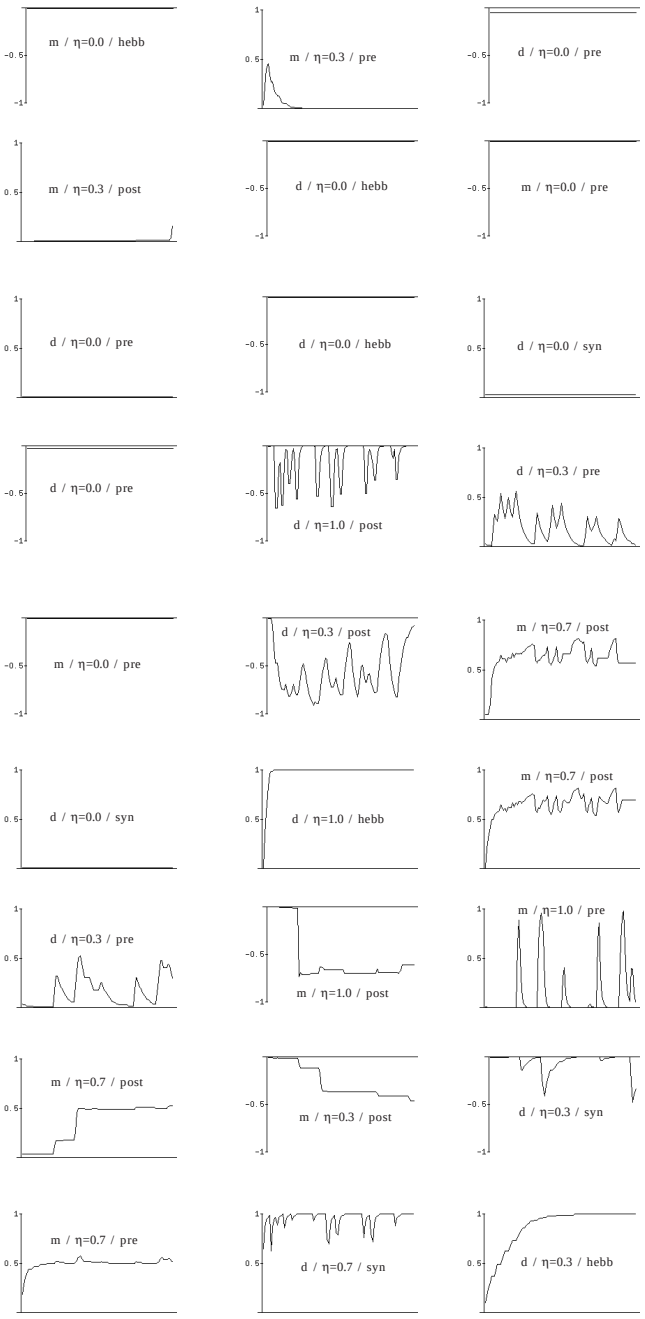

Figure 5: Synaptic strength recorded every 300 ms during the first 100 actions of the robot. The synapse layout is as depicted in Figure 2: rows represent input from the sensors (clockwise starting from the leftmost front sensor) and the hidden unit, columns represent input to the hidden unit, the left motor neuron and the right motor neuron. Details of the synapse characteristics are given in the following order: role (d = driving; m = modulatory), learning rate (η), learning rule (hebb = pure hebb; post = postsynaptic; pre = presynaptic; syn = synchrony). The sign of the y-axis indicates whether the synapse is excitatory or inhibitory.

neural network does not use the internal node to process the sensory information because most of the afferent synapses from the front sensors have the learning rate set to 0.0 (the presynaptic rule applied to the synapse from the second front sensor keeps its strength to 0.0 in normal conditions, i.e. when the robot follows a wall to its own right, but it can increase its efficacy when something is detected on the left). However, the internal node is maintained moderately active most of the time by excitatory connections from rear sensors and excitatory self-connection.

The behavioral preference for following walls on the right side is reflected by the null efficacy (for both wheels) of the afferent synapses from sensors on the left side: the only adjustable connection — from the leftmost sensor to the left wheel — uses a postsynaptic mechanism that enables only temporary excitation of the wheel and thus obstacle avoidance if something happens to be on the left side. Synapses from the internal node to the motor neurons are both driving and excitatory: given the constant level of activation of the internal node, these synapses provide a constant forward motion to both wheels. All the remaining synapses to the motor neuron that controls the right wheel are excitatory: that means that the right wheel will only move forward, its speed mainly generated by the internal node but modulated by the information coming from the two rightmost sensors. The excitatory driving synapse from the central sensor to the right wheel causes fast accelerations and decelerations which, combined with the opposite effect of the corresponding synapse to the other motor neuron, cause the backward rotations when something is frontally detected. Synapses connecting the three sensors on the front-right side to the left wheel are very important (they all convey driving signals): they are responsible for the wall-following behavior. The synapse corresponding to the rightmost sensor (nearly always highly active) quickly learns to transmit constant and high excitation to the left wheel. This excitation — that would otherwise lead the robot against the wall — is counterbalanced by the driving inhibitory inputs that come from the other two front-right sensors: the corresponding synapses display an oscillatory pattern that is responsible for the frequent small turns toward the wall. These turns are important because provide information about the wall curvature by allowing all the three right-front sensors to receive information on distance from the wall. The fast changing synapse from the front sensor provides rapid temporary inhibition to the left wheel when some object is frontally detected or the walls form a sharp convex angle smaller than 90 degrees; its pattern of strength change is in phase with that of the excitatory synapse from the same sensor to the right wheel: this causes the robot to stop and turn backward to re-adjust its trajectory. As the robot gradually adapts to the geometry of the environment,

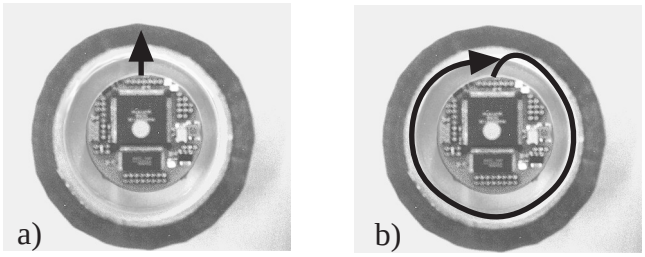

Figure 6: Sensory deprivation: a) the robot starts its life within the fence; b) the robot is put within the fence after the development of wall-following.

these turns become less frequent, which is reflected by the decreasing frequency and intensity of change of this synapse (and of the corresponding synapse to the other wheel).

4.2 Environment, Inherited Structures, and Learning

We have performed a number of tests on the same individual described above in order to investigate the influence of the training environment, and the relationships between inherited structure, early experiences, and learning. In a first test we have put the robot in an empty space and initialized the synaptic strengths with random values. After the initial short backward motion, the observed behavior was rather erratic, mostly composed of long forward motions and rapid turns in place to the left. If the robot was then put in the training environment, it would soon acquire the typical wall-following behavior. When put again in the empty space, it would start to alternate right turns (searching for the wall that suddenly disappeared) with sudden stops.

To better understand the role of early experience, we tried to produce learning in conditions of "sensory deprivation" by putting the robot within a small circular fence with an internal diameter of 75 mm (Figure 6). In this condition, all the sensors have roughly the same high activation, without regard to the action taken by the robot. When we put the robot at birth within the fence, it started to push in the frontal direction against the fence and it did not change behavior. Instead, when we put it within the fence after the usual learning phase in its normal environment, it began pushing frontally for one second or two, and then it started to turn slowly to the right still slightly pushing against the fence. The latter behavior, which is more varied and useful (it could help to find an opening –if there was one), could be achieved only if the phase of sensory-motor co-ordination had been properly completed in the normal environment.

In order to understand the relationship between the inherited neural structure and learning, we put the robot in the training environment, initialized its synaptic values to small random values, and disabled learning on all synapses. Typically, the robot would move straight and crash into the nearest wall without being able to recover, although in some cases it could manage to perform a left turn at the correct place. When we put this learning-impaired individual in an empty area, we observed that it performed long straight trajectories (approximately 40 cm) in the frontal direction interrupted by sharp right turns in place. Although this behavior does not allow proper navigation in a cluttered environment, nonetheless the trajectory roughly reflects the geometry of the environment where the robot has been evolved. These tests indicate that inherited structure does have a few basic and primitive skills (going forward and turning to the right) that have been shaped by the environment and the selection criterion. These inherited abilities, or "instincts", narrow the search space of learning providing a good starting point for a fast development of the ontogenetically acquired behavior.

5 Discussion

The genetic alphabet employed in these experiments was motivated by a set of computational considerations and neurophysiological findings. There is no special reason to believe that synaptic plasticity in biological systems can be explained in terms of a single learning mechanism; rather, individual synapses might modify their own strength according to different learning rules. The choice of a particular learning rule depends upon the types of receptors and transmitters found at the synaptic locus. Biologists have recently isolated a number of genes regulating the expression of NMDA receptors [9] which are thought to be the most likely mechanism responsible for Hebbian learning. On the other hand, neurophysiologists have provided evidence for the existence of a few types of Hebbian-like rules, where synapses are coincidence detectors that increase or decrease their own efficacy depending on the simultaneous activity level of the presynaptic and postsynaptic cells [10, 19, 16, 18]. Thus, it might be argued that the learning properties of synapses are specified in the genetic material, just like any other characteristics of the nervous system. Rather than coding chemical properties of synapses and simulating molecular dynamics, we simply coded four simple Hebbian rules for which there is neurophysiological evidence. These rules were modified by including a normalizing factor dependent on current synaptic efficacy that constrained maximum synaptic strength, as reported in [22], and could not change the sign of the synapse; this self-limiting mechanism had the computational advantage of avoiding the risk of saturating the activation function and thus reduced the search space of ontogenetic learning. Also the choice of encoding synaptic sign (which could not be in-

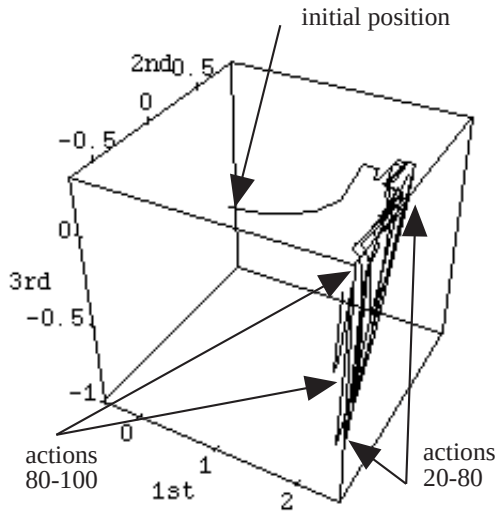

Figure 7: State-space representation of synapse dynamics during the first 100 actions plotted as trajectory within the space of the first three principal components. Arrows indicate the starting position and the range of oscillation between action sequences 20-80 and 80-100. Oscillations within the subspace of the third (smallest) component correspond to trajectory adjustments. Method: Sanger's network [15] for extracting the first three principal components of the input correlation matrix was trained to stability on the 27-component vectors corresponding to the synaptic activity recorded during the first 100 actions of the robot (visualized in Figure 5); after training, input vectors were presented again to the network and output unit activations were plotted.

verted by learning, as in biological nervous systems) was aimed at reducing the search space of ontogenetic learning and at putting more emphasis on inherited wiring of the neural structure. The choice of a two-component activation function combining driving and modulatory signals was motivated only by biological considerations; modulatory synapses, which are widespread in biological nervous systems and are modifiable (e.g., see [8]), are traditionally considered responsible for delivering contextual information [17, 3]. Here, we wanted to see how they could be exploited for sensory-motor control (note that when modulatory signals are absent, the activation function is equivalent to a standard sigmoid function).

Our results indicate that it is possible to evolve learning structures with emergent fast-adaptation properties. The evolved individual analyzed above displays interesting properties. The stable behavior acquired during life is regulated by continuously changing synapses which are *dynamically stable*. In the conventional view, synapses are relatively stable and slow components of the nervous system. Synaptic changes are identified with the learning of new skills or acquisition of new knowledge, while neuron activations are identified with the expression of behavior and of existing knowledge.[2] Typically, acquisition of a stable behavior in a static environment corresponds to stability (no further change) of individual synapses (e.g., see [7]). Such requirement is explicitly included into the objectives (least-mean-square error minimization, energy reduction, maximization of node mutual information, etc.) from which — both supervised and unsupervised — conventional learning algorithms are derived, but it is not included into the fitness function employed here, which is defined only in behavioral terms. The functioning of our system offers a complementary — but not necessarily alternative — explanation. Synapses are responsible for both learning and behavior regulation. Knowledge in the network is not expressed by a final stable state of the synaptic configuration, but rather by a dynamic equilibrium point in a n-dimensional state-space (where n is the number of synapses). Figure 7 plots the trajectory of synaptic change in the reduced state-space of the first three principal components of the recorded synaptic vectors during the first 100 actions of the individual displayed in Figure 4. During the first 6 actions the systems moves toward a subregion of the space for which there is no change in the first two principal components; residual variation along the slice of space corresponding to the third principal component corresponds to trajectory adjustments and is further reduced as the robot gradually tunes its path to the geometry of the environment.

In the case of the the neurocontroller analyzed above, most of the active synapses are excitatory (75%), but inhibition plays a key role in controlling some crucial aspects of the behavior, such as active sampling of the wall curvature and trajectory re-adjustment. Although there is not a preference for a particular learning rule, the evolutionary procedure has made sparse use of the plain Hebbian mechanism. This "choice" is quite reasonable because that rule does not allow reduction of synaptic efficacy and may thus hamper future adaptation. Where the Hebb rule has been employed, the links between the units (rightmost sensor to right and left wheel and internal node to right wheel) correspond to the establishment of a basic and "immutable" aspect of behavior, i.e. forward motion. The network employs the internal node as a sort of internal pattern generator[3] [2] that drives the organism forward even when there is not sensory stim-

[2] This view has been recently challenged by Yamauchi and Beer [21], who have evolved and analyzed continuous-time recurrent neural networks that give the external appearance of performing reinforcement learning while, in fact, these networks have fixed connection weights and use only internal node dynamics.

[3] The analogy should not be taken literally because the network dynamics and the use of wheels, rather than of legs, do not necessarily require the pattern of activity found in living organisms.

ulation available. It is interesting to note that internally generated forward motion is regulated by modulatory synapses from the two rightmost sensors to the right wheel: signals coming from these synapses contribute to adjust the distance of the robot from the wall.

The behavioral tests described in section 4.2 indicate that the structure of training environment plays a great role in shaping both the inherited neural structures and the properties of the ontogenetically developed behavior. Whatever the environment where the robot is placed, the neural network actively seeks an object and then performs a set of actions aimed at keeping it on its own right side. Its adaptation abilities are limited only to those variations that were encountered in the training environment (e.g., walls with different curvatures) and this is the reason for the development of a simple wall-following behavior. To this extent, such a plastic system evolved in stationary environments does not offer significant adaptation advantages w.r.t. systems with fixed synaptic weights. Our current research is focused on testing the procedure here described to changing environments.

A final remark concerns the robustness of the method employed. The relatively high variation in fitness values across runs (in the range ± 0.05 to ± 0.11) reflects a loose correspondence between the genotype specification and the phenotype performance: this means that a small change in the chromosome of the individual might result in drastic (and potentially maladaptive) changes at the phenotypic level. This is also reflected in the diversity –within a population and across runs– of the evolved neural structures. Although continuously changing synapses, dominance of excitatory vs. inhibitory synapses, and sparse use of plain Hebb learning are a common feature of the best 5-7 individuals in the last generation, all the remaining individuals display different structures and maladaptive behaviors. This variation can be partly explained by intrinsic variation of a not-yet converged population, but we think that it might be further reduced by employing more suitable building blocks (currently under investigation).

Appendix

Genetic algorithm parameters:

Population size	80
Generation number	50
Chromosome length	162 bits
Crossover probability	0.1
Mutation (expected probability of each bit being flipped)	0.2
Life length	80 actions
Action duration	300 ms

Acknowledgments

We would like to thank Edo Franzi and André Guignard for important work in the design of Khepera, and Randall Beer for useful comments on an early draft of this paper. Dario Floreano and Francesco Mondada have been partially supported by the Swiss National Research Foundation (project PNR23).

References

[1] D. H. Ackley and M. L. Littman. Interactions between learning and evolution. In C.G. Langton, J.D. Farmer, S. Rasmussen, and C. Taylor, editors, *Artificial Life II: Proceedings Volume of Santa Fe Conference*, volume XI. Addison Wesley: series of the Santa Fe Institute Studies in the Sciences of Complexities, 1991.

[2] F. Delcomyn. Neural basis of rhythmic behavior in animals. *Science*, 210:492–498, 1980.

[3] A. K. Engel, P. Koenig, A. K. Kreiter, T. B. Schillen, and W. Singer. Temporal coding in the visual cortex: new vistas on integration in the nervous system. *Trends in Neuroscience*, 15:218–226, 1992.

[4] D. Floreano and F. Mondada. Automatic creation of an autonomous agent: Genetic evolution of a neural-network driven robot. In D. Cliff, P. Husbands, J. Meyer, and S. W. Wilson, editors, *From Animals to Animats III: Proceedings of the Third International Conference on Simulation of Adaptive Behavior*. MIT Press-Bradford Books, Cambridge, MA, 1994.

[5] D. Floreano and F. Mondada. Evolution of homing navigation in a real mobile robot. *IEEE Transactions on Systems, Man, and Cybernetics-Part B*, 26:in press, 1996.

[6] D. E. Goldberg. *Genetic algorithms in search, optimization and machine learning*. Addison-Wesley, Reading, MA, 1989.

[7] J. Hertz, A. Krogh, and R. G. Palmer. *Introduction to the theory of neural computation*. Addison-Wesley, Reedwood City, CA, 1991.

[8] J. A. Hirsch and C. D. Gilbert. Long-term changes in synaptic strength along specific intrinsic pathways in the cat visual cortex. *Journal of Physiology*, 461:247–262, 1993.

[9] M. Hollman and S. Heinemann. Cloned glutamate receptors. *Annual Review of Neuroscience*, 17:31–108, 1993.

[10] S.R. Kelso, A.H. Ganong, and T.H. Brown. Hebbian synapses in hippocampus. *Proceedings of the National Academy of Sciences USA*, 83:5326–5330, 1986.

[11] D. J. McFarland. *The Oxford Companion to Animal Behaviour*. Oxford University Press, Oxford, 1981.

[12] F. Mondada, E. Franzi, and P. Ienne. Mobile robot miniaturization: A tool for investigation in control algorithms. In *Proceedings of the Third International Symposium on Experimental Robotics*, Kyoto, Japan, 1993.

[13] D. Parisi, F. Cecconi, and S. Nolfi. Econets: Neural networks that learn in an environment. *Network*, 1:149–168, 1990.

[14] W. A. Phillips, J. Kay, and D. Smyth. The discovery of structure by multi-stream networks of local processors with contextual guidance. *Network*, 6:225–246, 1995.

[15] T. D. Sanger. Optimal unsupervised learning in a single-layer feedforward neural network. *Neural Networks*, 2:459–473, 1989.

[16] W. Singer. Activity-dependant self-organisation of synaptic connections as a substrate of learning. In J. P. Changeux and M. Konishi, editors, *The Neural and Molecular Bases of Learning.* Wiley, London, 1987.

[17] W. Singer. Search for coherence: a basic principle of cortical self-organization. *Concepts in Neuroscience*, 1:1–26, 1990.

[18] P. K. Stanton and T. J. Sejnowski. Associative long-term depression in the hippocampus induced by hebbian covariance. *Nature*, 339:215–218, 1989.

[19] G.S. Stent. A physiological mechanism for hebb's postulate of learning. *Proceedings of the National Academy of Sciences USA*, 70:997–1001, 1973.

[20] D. Willshaw and P. Dayan. Optimal plasticity from matrix memories: What goes up must come down. *Neural Computation*, 2:85–93, 1990.

[21] B. Yamauchi and R. D. Beer. Sequential behavior and learning in evolved dynamical neural networks. *Adaptive Behavior*, 2(3):219–246, 1995.

[22] X-D. Yang and D. S. Faber. Initial synaptic efficacy influences induction and expression of long-term changes in trasmission. *Proceedings of the National Academy of Science USA*, 88:4299–4303, 1991.